

THE Pa α LUMINOSITY FUNCTION OF H II REGIONS IN NEARBY GALAXIES FROM *HST*/NICMOS*

GUILIN LIU^{1,2}, DANIELA CALZETTI¹, ROBERT C. KENNICUTT, JR.³, EVA SCHINNERER⁴, YOSHIKI SOFUE⁵,
 SHINYA KOMUGI⁶, FUMI EGUSA⁷, AND NICHOLAS Z. SCOVILLE⁸

¹ Astronomy Department, University of Massachusetts, Amherst, MA 01003, USA

² Center for Astrophysical Sciences, Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA; liu@pha.jhu.edu

³ Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK

⁴ MPI for Astronomy, Königstuhl 17, D-69117 Heidelberg, Germany

⁵ Department of Physics, Meisei University, 2-1-1 Hodokubo, Hino, Tokyo 191-8506, Japan

⁶ Joint ALMA Office, Alonso de Córdova 3107, Vitacura, Santiago 763-0355, Chile

⁷ Institute of Space and Aeronautical Science, Japan Aerospace Exploration Agency, 3-1-1 Yoshino-dai, Chuo-ku, Sagami-hara, Kanagawa 252-5210, Japan

⁸ California Institute of Technology, MC 249-17, Pasadena, CA 91125, USA

Received 2013 April 9; accepted 2013 May 21; published 2013 July 2

ABSTRACT

The H II region luminosity function (LF) is an important tool for deriving the birthrates and mass distribution of OB associations and is an excellent tracer of the newly formed massive stars and associations. To date, extensive work (predominantly in H α) has been done from the ground, which is hindered by dust extinction and the severe blending of adjacent (spatially or in projection) H II regions. Reliably measuring the properties of H II regions requires a linear resolution <40 pc, but analyses satisfying this requirement have been done only in a handful of galaxies, so far. As the first space-based work using a galaxy sample, we have selected 12 galaxies from our *HST*/NICMOS Pa α survey and studied the LF and size distribution of H II regions both in individual galaxies and cumulatively, using a virtually extinction-free tracer of the ionizing photon rate. The high angular resolution and low sensitivity to diffuse emission of NICMOS also offer an advantage over ground-based imaging by enabling a higher degree of de-blending of the H II regions. We do not confirm the broken power-law LFs found in ground-based studies. Instead, we find that the LFs, both individual and co-added, follow a single power law $dN(L)/d \ln L \propto L^{-1}$, are consistent with the mass function of star clusters in nearby galaxies, and are in agreement with the results of the existing analyses with *Hubble Space Telescope* (*HST*) data. The individual and co-added size distributions of H II regions are both roughly consistent with $dN(D)/d \ln D \propto D^{-3}$, but the power-law scaling is probably contaminated by blended regions or complexes.

Key words: galaxies: individual – H II regions

Online-only material: machine-readable table

1. INTRODUCTION

As a well-established tracer of the newly formed massive stars and star associations and clusters, the distribution of H II regions has sparked a long interest since the pioneering investigations in nearby spiral galaxies, as reviewed by Kennicutt (1992). Extensive statistical analyses of galactic H II regions and complexes have been undertaken during the recent couple of decades, predominantly focusing on the luminosity functions (LFs), an important tool for deriving the birthrates and mass distribution of OB associations. These studies, mostly based on ground-based imaging in the hydrogen H α emission line, have yielded a collection of remarkable results. Notably, Kennicutt et al. (1989) analyzed a sample of 30 galaxies with Hubble type ranging from Sb to Irr and found the locus of differential LFs to follow a power law expressed as $dN/dL \propto L^{-2 \pm 0.5}$ for luminosities $L_{H\alpha} \gtrsim 10^{37}$ erg s $^{-1}$. They also found that a subsample ($\sim 20\%$ of the total) of the galaxies is better described by a double power-law LF, which they termed “type II,” with a break in the slope for luminosities $L_{H\alpha} < 10^{38.7}$ erg s $^{-1}$. The H α analyses of H II region LFs were later extended to early-type (Sa)

spirals (Caldwell et al. 1991), H I-selected galaxies (Helmboldt et al. 2005), and low surface brightness galaxies (Helmboldt et al. 2009).

Two conventions for the representation of the LFs exist in the literature, one for number counts binned linearly in luminosity and the other binned logarithmically. We adopt the logarithmic formalism, $dN/d \ln L \propto L^\alpha$, equivalent to $dN/dL \propto L^{\alpha-1}$ (note α is generally negative). The existence of the slope break at an H α luminosity of $L_{br} \sim 10^{38.5-38.7}$ erg s $^{-1}$ is confirmed in the LF of M101 by Scowen et al. (1992), M51 by Rand (1992), eight other galaxies by Rozas et al. (1996, 1999, 2000), and more recently, in the co-added LF derived from the 18,000 H II regions in 53 galaxies by Bradley et al. (2006). This high-luminosity break is hence considered robust and termed the Strömgren luminosity by Beckman et al. (2000), in accordance with their hypothesis that the break results from a transition from ionization-bounded to density-bounded H II regions. Conversely, Oey & Clarke (1998) suggest that the luminosity limit $L_{H\alpha} \sim 10^{38.5}$ erg s $^{-1}$ is linked to the regions transitioning from stochastic sampling of the stellar initial mass function (IMF) to full sampling. The ground-based H α LFs generally show a “type II” profile when a steep slope at the bright end ($\alpha_{br+} \sim -1$) is connected to a flatter one in the less luminous regime ($\alpha_{br-} \sim -(0.3-0.5)$; see Section 3.1 for Pa α LFs of individual galaxies). Here and throughout this section, H II regions brighter (fainter) than L_{br} are called *super-break*

* Based on observations taken with the NASA/ESA *Hubble Space Telescope* obtained at the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555. These observations are associated with program GO-11080.

(*sub-break*), and the quantities that describe their behavior are marked with a subscript “br+” (“br−”).

The hypothesis that the break in the slope has a physical nature rather than being an observational artifact has been under suspicion for a long time, especially when considering that the resolution of ground-based images, dominated by seeing, ranges from $0''.8$ to $3''-4''$, corresponding to up to ~ 400 pc at a distance of 20 Mpc. The ground-based work is potentially hindered by blending of adjacent (spatially or in projection) H II regions. This results in an artificially pumped-up bright end of the number distribution together with a depleted faint end. As shown by Pleuss et al. (2000), measuring the properties of H II regions reliably requires a linear resolution of ~ 40 pc or better, translating to an angular resolution of $\lesssim 0''.4$ for a galaxy located 20 Mpc away. The *Hubble Space Telescope* (*HST*), which has a $<0''.3$ resolution (actual value depending on the instrument), is thus an ideal facility to tackle the H II region statistics.

The high angular resolution of the *HST* data motivated Scoville et al. (2001) and Pleuss et al. (2000) to scrutinize the previous conclusions drawn for M51 and M101, respectively, using *HST* H α images. Interestingly, neither of the two galaxies shows evidence for the broken power-law H α LF obtained in ground-based studies. Both galaxies exhibit a single power-law LF characterized by $\alpha \simeq -1$, or equivalently, $\alpha_{\text{br-}} \simeq -1$, which is much steeper than previously determined $\alpha_{\text{br-}}$ values. These values of $\alpha_{\text{br-}}$ are consistent with results on the luminosity and mass functions of star clusters (e.g., a recent analysis of *HST* WFC3 data of M83 and M51; Chandar et al. 2010, 2011). In addition, both galaxies contain only a small amount ($\lesssim 1\%$ of the total) of super-break H II regions. This result indicates that super-break H II regions may not exist after all, but are simply blends of multiple regions and associated diffuse gas. If this is correct, the break at L_{br} is actually where the power-law scaling truncates. The discrepancy between these *HST* analyses of two galaxies and the more extensive previous ground-based work brings urgency to a more systematic investigation using the high angular resolution data of the *HST*. Although the more recent *HST* study by Lee et al. (2011) finds evidence for the existence of a break luminosity in the H II LF of M51, the break occurs at a luminosity fainter than the type II LF break by a factor of ~ 30 ($L_{\text{H}\alpha} = 10^{37.1}$ erg s $^{-1}$) and is thus likely to have a different origin. Meanwhile, the upper cutoff of the H II LF of M31 is below the type II LF break luminosity, disqualifying this galaxy for testing the above discrepancy (Azimlu et al. 2011).

Far less investigated than the impact of blending is the effect of dust attenuation on the derivation of the LFs. Variable dust attenuation across galactic disks induces scatter on the measurements. Furthermore, since brighter H II regions are generally associated with larger extinctions (e.g., Calzetti et al. 2007), the presence of dust can hinder accurate derivations of H II LFs and also the verification of the presence of a luminosity truncation, as discussed in Scoville et al. (2001). Less susceptible to dust attenuation, near-IR tracers of current star formation activity, such as the Paschen or Brackett emission lines, can be used as good alternatives to H α and other optical tracers. Such measurements are only available for very limited samples, but an *HST*/NICMOS Pa α survey of the central $\sim 1'$ field of 84 nearby galaxies has been recently completed (PI: Daniela Calzetti; GO: 11080). The Pa α hydrogen line emission, with a central rest-frame wavelength of $1.8756 \mu\text{m}$, suffers much less from dust attenuation than H α . The visual extinctions of individual H II regions in the almost face-on spirals like M51

can reach values as high as $A_V \simeq 6$ mag (Scoville et al. 2001), implying that the H α line is dimmed by two orders of magnitude, but Pa α only a factor of ~ 2 . In general, we expect a 20% dimming factor for Pa α to be more typical in nearby disk galaxies (Calzetti et al. 2007).

In parallel to the LF investigations, the size distribution of the H II regions, physically related to the LF, has also been studied at length (e.g., Hodge et al. 1989; Knapen 1998; Petit 1998; Youngblood & Hunter 1999; Hodge et al. 1999). More recently, Oey et al. (2003) compiled and refitted the data in the literature and suggest that the differential nebular size distribution follows a power law of the form $N(D)dD \propto D^{-4}dD$, with flattening at diameters below ~ 260 pc. Although the regions that enter in the power-law distribution are clearly blends of multiple smaller objects, their investigation provides important clues to the clustering properties of H II regions.

HST is relatively insensitive to the low surface brightness emission from the diffuse ionized gas due to its intrinsic high angular resolution (Kennicutt et al. 2007). Thus, our *HST* Pa α images are only sensitive to the compact emission from associations of relatively massive ($\gtrsim 30 M_{\odot}$), young ($\lesssim 8-10$ Myr) stars. This feature, in addition to its high angular resolution, further makes characterizing H II regions more reliable, as the associated diffuse gas also plays an important role in preventing H II regions from being de-blended (Pleuss et al. 2000).

We select 12 spirals out of the 84 targets in our *HST*/NICMOS program to construct our sample so that every member contains at least 50 well-detected H II regions (see Section 2.1 for further details). Our sample ranges from Sb to Sd in morphological type (see Table 1); these are later types than the sample as a whole, and are a result of the bias introduced by the requirement that each galaxy contains an appreciable number of H II regions. This is our best sample to infer the intrinsic (i.e., minimally affected by blending and dust attenuation) LF of H II regions with high accuracy. This work is the first systematic investigation of the H II region statistics utilizing the *HST* Pa α images of a significant galaxy sample.

The paper is organized as follows: Section 2 describes the basic properties of our galaxy sample and provides some details on the data reduction and the identification and photometry of H II regions; Section 3 presents our analysis of H II LF and size distribution in individual galaxies; Section 4 is devoted to the analysis of co-added H II region distributions by combining the whole sample, and its comparison to previous ground-based studies in H α ; the results are discussed in Section 5 and summarized in Section 6.

2. DATA SAMPLE

2.1. Sample Description

The galaxies analyzed in this paper are selected from a recent *HST*/NICMOS survey of the central $\sim 1'$ field of 84 nearby galaxies evenly distributed across all Hubble and bar types (*HST*/GO-11080). Targeting the redshifted hydrogen Pa α emission line using the NIC3 camera, these observations were completed in 2008 February.

In order to optimize our analysis, we downselect a subsample from the original list of 84 galaxies according to the following criteria.

1. Highly inclined spirals ($i > 70^\circ$) are excluded to avoid severe blending of overlapping (in projection) H II regions.
2. Galaxies with heliocentric recession velocities higher than 1600 km s^{-1} are rejected. For larger redshifts, the

Table 1
Characteristics of the Sample Galaxies and Their *HST* Pa α Data

Galaxy (1)	Morphology (2)	T (3)	i (4)	Distance (5)	Referene (6)	FOV (7)	Resolution (8)	AGN Type (9)
NGC 278	SAB(rs)b	2.9	12.8	11.8	1	2.9	15	
NGC 972	Sab	2.0	65.9	15.5	2	3.8	20	
NGC 1068	(R)SA(rs)b	3.0	21.1	10.1	2	2.5	13	Seyfert 1.8 ^a
NGC 1097	SB(s)b-LINER b	3.3	37.0	14.2	2	3.5	18	Seyfert 1 ^b
NGC 3310	SAB(r)bc pec-H II	4.0	31.2	17.5	3	4.3	22	
NGC 3621	SA(s)d	6.9	65.6	7.11	4	1.8	9	
NGC 3810	SA(rs)c	5.2	48.2	15.3	2	3.8	19	
NGC 3982	SAB(r)b?	3.2	29.9	21.2	5	5.2	27	Seyfert 1.9 ^a
NGC 4041	SA(rs)bc?	4.0	22.0	22.7	1	5.6	29	
NGC 4088	SAB(rs)bc	4.7	69.4	14.3	2	3.5	18	
NGC 4414	SA(rs)c?	5.2	54.0	21.4	6	5.3	27	trans. obj. 2 ^a
NGC 5248	SAB(rs)bc	4.0	56.4	12.7	2	3.1	16	

Notes. Column 1: Galaxy name; Column 2: morphology from the NASA/IPAC Extragalactic Database (NED, <http://ned.ipac.caltech.edu>); Column 3: numerical Hubble type (as defined in de Vaucouleurs et al. 1991) from the HyperLeda database (Paturel et al. 2003, <http://leda.univ-lyon1.fr>); Column 4: inclination ($^{\circ}$) from the HyperLeda database; Column 5: distance (Mpc) from the literature listed in Column 6; Column 6: references for distances—1, Tully 1988; 2, Tully et al. 2009; 3, Terry et al. 2002; 4, Rizzi et al. 2007; 5, Ngeow & Kanbur 2006; 6, Saha et al. 2006; Column 7: linear scale of field of view (kpc); Column 8: linear resolution (pc); Column 9: AGN type, if it exists—(a) from Ho et al. (1997); (b) from NED.

throughput of F187N filter is low, leading to a large transmittance correction which yields uncertain Pa α fluxes.

3. To ensure the fidelity of the statistical analysis, a galaxy is rejected if it contains fewer than 50 H II regions detected with signal-to-noise ratio $S/N \geq 3$ within the field of view (FOV) of the NIC3 camera. If the galaxy hosts a central active galactic nucleus (AGN), these H II regions need to be >200 pc from the galaxy center to avoid contamination from it. As a reference, our observations cover the central 5 kpc or less of each galaxy.
4. Only spirals located within 23 Mpc are considered, implying a worst spatial resolution of 29 pc (for the $0''.26$ resolution of our NIC3 images; see next section), well below the 40 pc maximum size suggested by Pleuss et al. (2000) for reliably measuring H II region properties.

The above selection procedure filters out 6/7 of the original sample, leaving 12 spiral galaxies (Figure 1). The Hubble type of these selected galaxies ranges from Sab to Sd, which, as mentioned in the Introduction, is due to the third selection criterion that biases the sample toward later type spirals. The basic properties of our sample galaxies are summarized in Table 1.

2.2. Observations and Data Reduction

The 12 sample galaxies were observed between 2007 February 13 and 2008 February 3 with the NICMOS/NIC3 camera on board the *HST*. NIC3 images have a $51'' \times 51''$ FOV and a pixel scale of $0''.2$ with an undersampled point-spread function (PSF). In this survey, each observation consists of images taken in two narrowband filters: one centered on the Pa α hydrogen recombination line ($1.87 \mu\text{m}$, F187N) and the other on the adjacent narrowband continuum (F190N). Each set of exposures was made with a six-position small ($<1''$ step) dither to maximize recovery of spatial information and rejection of cosmic rays and their persistence. Exposures of 160 and 192 s per dither position in the narrowband filters targeting the line emission and the adjacent continuum, respectively, reach a 1σ detection limit of $4.5 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$ in the continuum-subtracted Pa α images. Assuming the IMF given by Kroupa (2001) and applying the relation between star formation rate (SFR) and Pa α

luminosity (Calzetti 2007), we find the limiting SFR per unit area to be $\sim 0.0015 M_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$.

The data reduction steps consist of removing the NICMOS *Pedestal* effect with the *pedsub* task in the STSDAS package of IRAF. The offsets between the dither images are then found using the *crossdriz* and *shiftfind* tasks by computing the cross-correlation functions of the individual maps. After that, we apply the data quality files to create a static pixel mask for each dither image before these 256×256 pixel maps are drizzled onto a larger (1024×1024) and finer ($0''.1 \text{ pixel}^{-1}$) grid frame using the *drizzle* package. A single median image is then made by combining all the drizzled images using the offsets determined in the previous step. In order to reject cosmic rays, the median image is “blotted” or reverse-drizzled back to the original dimensions of each original image, and also shifted. By creating the derivative images of the outputs with the *deriv* task, the value of each pixel essentially represents the largest gradient across that pixel. A comparison between the median and blotted images enables us to produce a cosmic-ray mask using the *driz_cr* task. The static mask file is then multiplied with each individual bad pixel mask file. Equipped with these “master” mask files, we drizzled each input image onto a single output image, applying the shifts and the mask files previously created. More details are available in the *HST Dither Handbook*,⁹ which we followed to derive our final images.

After applying all the steps above on both the F187N and F190N images, we aligned each filter pair to each other, scaled the F190N image to the F187N using the filter transmission curves, and subtracted the former from the latter to produce a stellar-continuum-free line emission image. We note that the line emission maps suffer slightly from residual shading, by exhibiting a pair of vertical banding features. We remove the artifact column by column, by calculating the median value of each column of the image and subtracting it. This strategy works remarkably well in the case of sparse emission blobs (the typical case in our images). The global image background is also removed by this subtraction. The transmittance correction for the line emission images is finally performed, adopting the recession velocities from the literature and the F187N

⁹ http://www.stsci.edu/hst/HST_overview/documents

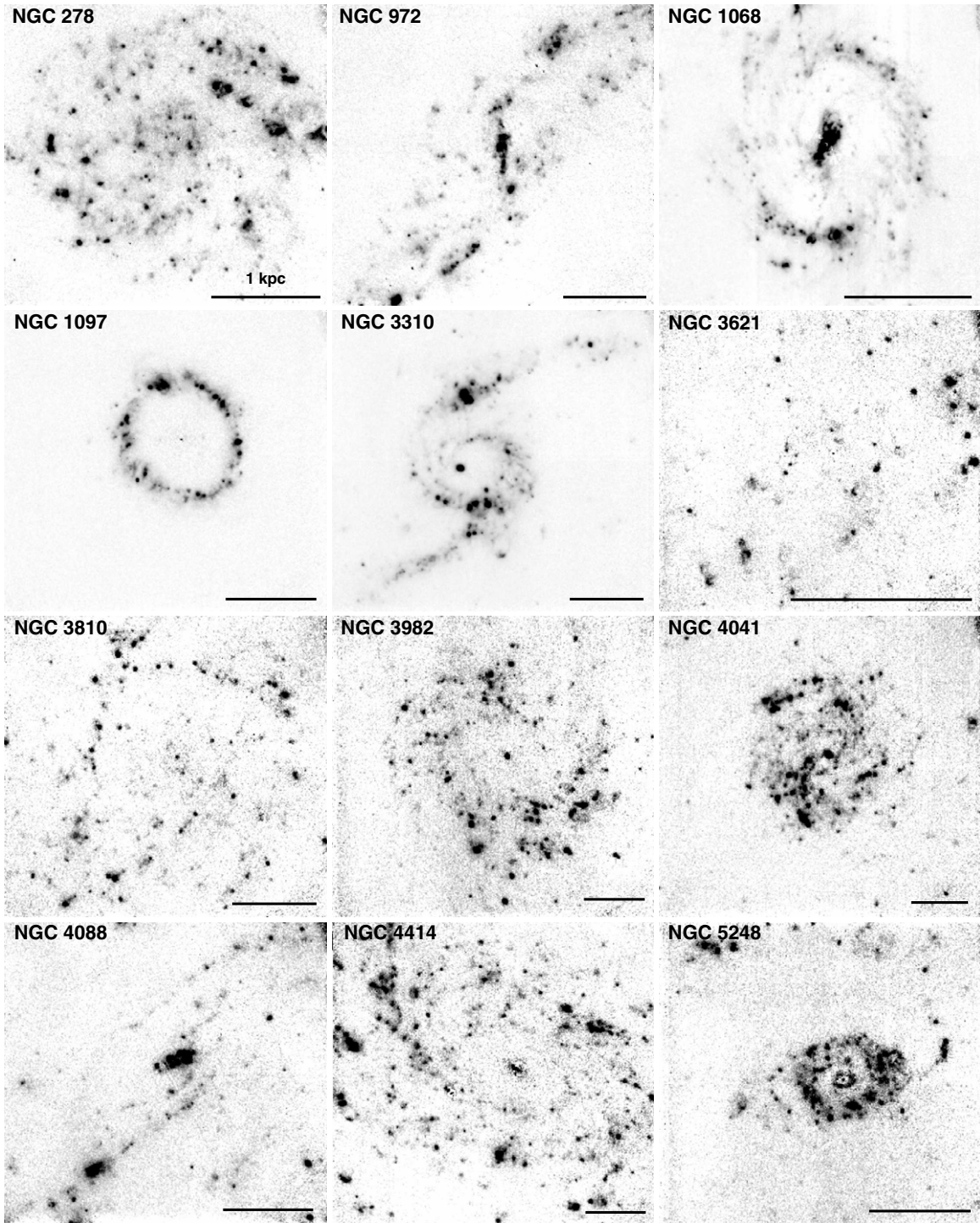


Figure 1. *HST*/NICMOS $\text{Pa}\alpha$ images of the 12 star-forming spiral galaxies that form our sample. The underlying stellar continuum has been subtracted from the line emission images. The data have an angular resolution of $0''.26$ and a field of view of $51''$ for each target. The bar at the bottom right corner of each panel shows the linear scale of 1 kpc.

throughput curve from the *HST* instrument handbook. The final continuum-subtracted $\text{Pa}\alpha$ images have PSFs of $0''.26$ (FWHM). The $\text{Pa}\alpha$ data of two galaxies from this *HST* campaign analyzed by Hsieh et al. (2011, NGC 1097) and van der Laan et al. (2013, NGC 6951) were reduced following an identical procedure.

2.3. H II Region Identification and Photometry

The identification, photometry, and cataloging of H II regions in the *HST* $\text{Pa}\alpha$ images are carried out using the IDL program *HIIphot*. This software, developed by Thilker et al. (2000), is

specifically designed for automated photometry and statistical analysis of H II regions on $\text{H}\alpha$ images of nearby galaxies, and is directly applicable to our $\text{Pa}\alpha$ data. Its capability to deal with crowded fields and adapt to irregular source morphologies results in accurate photometric characterization of H II regions.

The *HIIphot* algorithm can be summarized as follows. On a pure hydrogen emission line image of a galaxy, it identifies H II regions by smoothing the image with kernels of different sizes, utilizes object recognition techniques to fit models (“seeds”) to peaks that are identified via a minimum assigned S/N, and then

Table 2
Catalog of the 1907 H II Regions Detected with $S/N \geq 3$
in Our Sample Galaxies

Galaxy (1)	ID (2)	R.A. (3)	Decl. (4)	$\log L_{\text{Pa}\alpha}$ (5)	D (6)	S/N (7)
NGC 278	1	00:52:06.7	+47:33:00.2	36.79	33.4	3.4
NGC 278	2	00:52:06.7	+47:33:01.2	37.13	39.8	5.3
NGC 278	3	00:52:06.7	+47:33:02.2	37.07	30.7	5.3
NGC 278	4	00:52:06.7	+47:33:03.6	37.16	46.3	5.7
NGC 278	5	00:52:06.7	+47:32:57.0	36.85	43.4	3.5
NGC 278	6	00:52:06.7	+47:33:01.3	38.13	112.0	17.1
NGC 278	7	00:52:06.6	+47:33:05.4	37.28	62.1	6.9
NGC 278	8	00:52:06.6	+47:33:03.8	36.78	33.4	3.5
NGC 278	9	00:52:06.6	+47:33:03.1	37.43	53.2	7.8
NGC 278	10	00:52:06.6	+47:33:04.6	36.92	41.4	4.3
NGC 278	11	00:52:06.6	+47:33:02.7	37.84	67.7	13.0
NGC 278	12	00:52:06.6	+47:33:00.7	37.02	37.0	4.5
NGC 278	13	00:52:06.6	+47:33:03.6	37.13	43.4	5.4
NGC 278	14	00:52:06.6	+47:33:01.1	37.35	43.4	7.0
NGC 278	15	00:52:06.5	+47:33:02.2	37.84	68.0	12.8

Notes. Column 1: Galaxy name; Column 2: identification number of the H II region, starting from one for each galaxy; Column 3: right ascension of the identified region (hh:mm:ss.s); Column 4: declination of the region (dd:mm:ss.s); Column 5: $\text{Pa}\alpha$ luminosity of the region (erg s^{-1}) in logarithmic form. Uncertainties on $L_{\text{Pa}\alpha}$ can be estimated as explained in the text (Section 2.3); Column 6: equivalent diameter which matches the area of a circle $\pi D^2/4$ to that of the region (pc); Column 7: signal-to-noise ratio of the region.

(This table is available in its entirety in a machine-readable form in the online journal. A portion is shown here for guidance regarding its form and content.)

assigns pixels to identified H II regions. After that, an iterative growing procedure allows for departures from these “seeds.” The boundaries are built up at successively fainter isophotal levels that keep growing until either boundaries of other regions are encountered, or a pre-set limiting gradient of surface brightness is reached, implying its arrival at the background level. Two-dimensional interpolations are then performed by fitting the background pixels around the designated H II regions to estimate the local background behind the source. The final information on the detected sources is recorded in a file for further analysis.

The astrometric and photometric information of all the 1907 H II regions detected with $S/N \geq 3$ is tabulated in Table 2. Note that the calibration of our *HST* data is accurate within $\sim 5\%$, the total uncertainty introduced by instrumental artifacts and continuum subtraction is $\sim 10\%$, and thus the relative uncertainty of the derived $\text{Pa}\alpha$ luminosity can be estimated as $(0.05^2 + 0.10^2 + (S/N)^{-2})^{1/2}$, which is about 35% for an H II region detected with $S/N = 3$.

In agreement with the range of electron temperatures generally derived for H II regions in metal-rich spiral galaxies (Bresolin et al. 2004; Li et al. 2013), throughout the rest of this section we adopt $T_e \sim 5000\text{--}10,000$ K. We also adopt $n_e = 100 \text{ cm}^{-3}$, implying an intrinsic range $H\alpha/\text{Pa}\alpha = 7.4\text{--}8.5$, i.e., a mean value of 8 and a dispersion $\lesssim 10\%$ (Osterbrock & Ferland 2006). Although our measured n_e are significantly smaller than those adopted here (see Section 3.3), we recall that the hydrogen line ratio has a weak dependence on the adopted n_e .

Since we do not have dust extinction maps for our sample, we use H II regions’ extinction distributions from the literature as our reference and assume that our galaxies behave in a similar way as other galaxies. The dust extinction distribution for H II regions in M51, measured by Scoville et al. (2001) using *HST*

data both in $H\alpha$ and $\text{Pa}\alpha$, covers the range $A_V = 0\text{--}6$ mag, with $\sim 80\%$ of the regions having A_V values between 2 and 4 and a mean value of 3.1. Adopting the Galactic extinction curve of Cardelli et al. (1989), this extinction translates into an observed $H\alpha/\text{Pa}\alpha$ line ratio between 0.72 and 2.47, with a mean value of 1.25. Calzetti et al. (2007) measure the dust extinction distribution in 33 nearby galaxies, using *HST* $\text{Pa}\alpha$ and ground-based $H\alpha$. Although the low resolution of the ground $H\alpha$ images may reduce the measured extinction, they still obtain a median $A_V = 2.2$ mag, which translates into an observed $H\alpha/\text{Pa}\alpha$ line ratio of 2.2. For practical purposes, these measurements mean that for an *observed* break luminosity in $H\alpha$, $L_{H\alpha, \text{br}} \sim 10^{38.6} \text{ erg s}^{-1}$, the corresponding $L_{\text{Pa}\alpha, \text{br}} \sim 10^{38.3\text{--}38.5} \text{ erg s}^{-1}$, only 0.1–0.3 dex lower than the unextinguished break luminosity in $\text{Pa}\alpha$.

3. $\text{Pa}\alpha$ LUMINOSITY FUNCTION AND SIZE DISTRIBUTION

3.1. Individual Galaxies: $\text{Pa}\alpha$ Luminosity Functions

Similar to the $H\alpha$ H II region LF, the bright end of which is often represented by a truncated power law, we express the $\text{Pa}\alpha$ LF as, following McKee & Williams (1997) and Scoville et al. (2001),

$$dN(L_{\text{Pa}\alpha})/d \ln L_{\text{Pa}\alpha} = N_{\text{up}}(L_{\text{Pa}\alpha}/L_{\text{up}})^{\alpha}, \quad (1)$$

where the power index $\alpha < 0$, L_{up} is the brightest H II region, and N_{up} is roughly the number of regions with luminosities $0.5 < L_{\text{Pa}\alpha}/L_{\text{up}} < 1$ in case $\alpha \approx -1$. Further discussion on the physical interpretation of the parameters is deferred to the original papers.

The results of fitting Equation (1) to the observed $\text{Pa}\alpha$ LF in our sample galaxies are presented in Figure 2. In this figure, all the H II regions with $S/N \geq 3$ are plotted, but only the regions above the completeness limit for luminosity are used for the power-law fits. $S/N = 5$ represents the completeness limit recommended by Thilker et al. (2000) and is converted to a luminosity limit in Figure 2 (vertical dotted lines) by performing power-law fits to the well-defined luminosity vs. S/N correlation. One should note that this completeness limit of H II region identification and characterization is, unsurprisingly, much higher than the detection limit of the *HST* imaging quoted in the observations and data reduction section (Section 2.2). For the whole sample, we find a mean value for the LF power index $\langle \alpha \rangle = -1.05$, and a median value $\tilde{\alpha} = -0.98 \pm 0.19$.

As mentioned in the Introduction, some galaxies are characterized by LFs described by a broken, but connected, power law, with a luminosity at the break $L_{H\alpha, \text{br}} \sim 10^{38.6} \text{ erg s}^{-1}$. To provide some perspective on this value, we recall that a single O-type star with mass $120 M_{\odot}$ produces an ionizing photon rate of $\sim 10^{50.1} \text{ s}^{-1}$ (Schaerer & de Koter 1997; Oey & Clarke 1998), which corresponds to an intrinsic $L_{H\alpha} \sim 10^{38.2} \text{ erg s}^{-1}$ (Leitherer & Heckman 1995), and a dust-attenuated $H\alpha$ luminosity of $\sim 10^{37.2\text{--}37.5} \text{ erg s}^{-1}$, for the dust extinction values discussed in Section 2.3. The observed $H\alpha$ luminosity for a single O-star is thus well below the value of the break luminosity mentioned above and close to the completeness limits of H II region identification shown in Figure 2.

The broken power-law LF does not seem to be present for the H II LF of our sample galaxies. All the 12 galaxies exhibit $\text{Pa}\alpha$ LFs that follow a single power law with $\alpha \simeq -1$ above the completeness limit, with no apparent break around $10^{38.2\text{--}38.4} \text{ erg s}^{-1}$

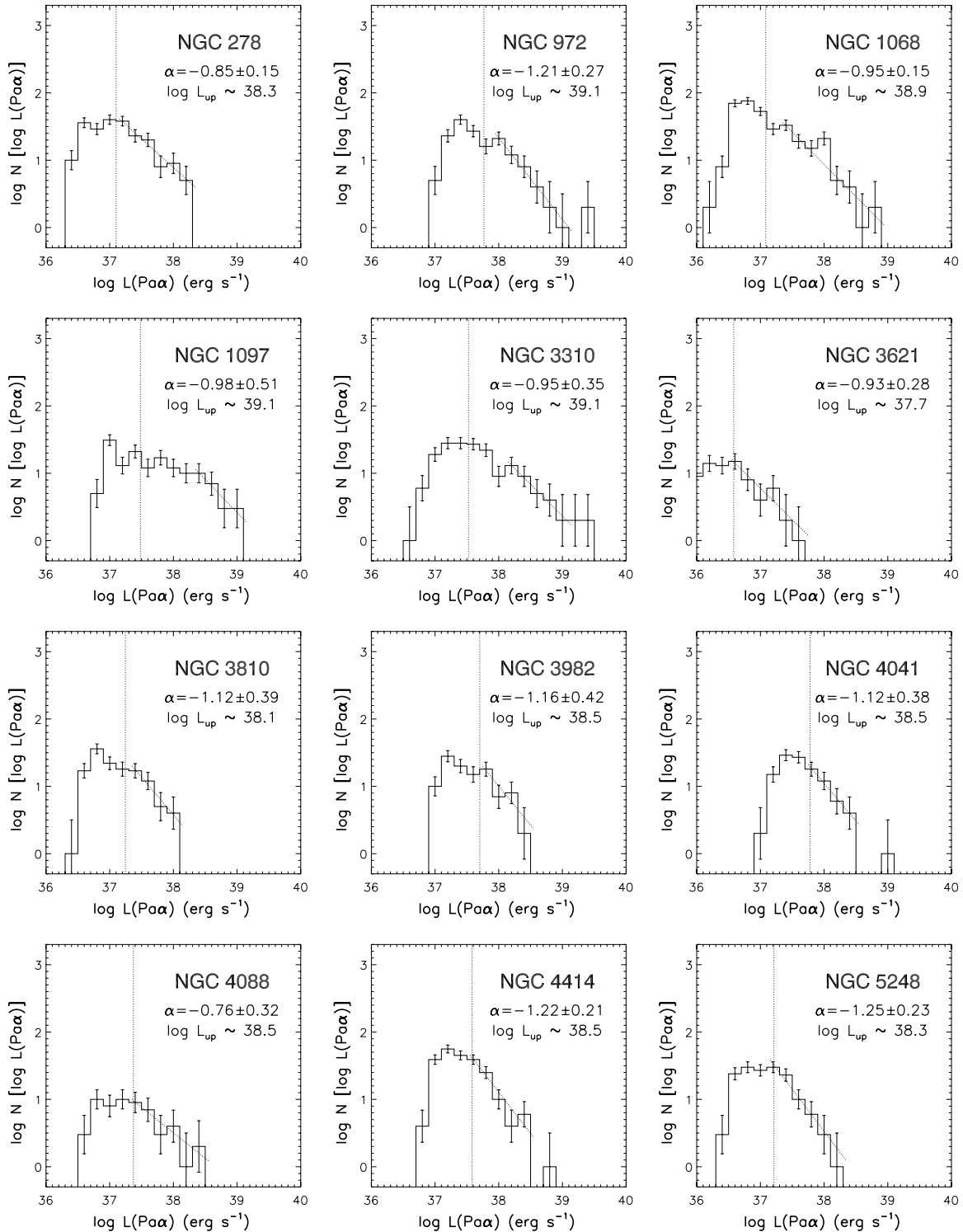


Figure 2. $\text{Pa}\alpha$ H II region luminosity functions of the 12 sample galaxies. Each bin containing N regions is weighted by \sqrt{N} when the power-law fits are performed. Plotted are all regions with $S/N \geq 3$, but the power-law fitting is performed on the bright end of regions above the $S/N = 5$ completeness limit depicted by the vertical dotted lines.

(Figure 2), consistent with the results for Galactic radio H II regions, which show a single power-law index of -1.3 (Smith & Kennicutt 1989) or -1.0 (McKee & Williams 1997). This conclusion also coincides with the result found from the $\text{H}\alpha$ *HST*/WFPC2 imaging of M51 ($\alpha = -1.01 \pm 0.04$, no break; Scoville et al. 2001), although it is steeper than what is found for M101 ($\alpha = -0.74 \pm 0.08$, no break; Pleuss et al. 2000).

Figure 2 also shows the location of a high-luminosity truncation, L_{up} , where the power-law bright end is reached because of small number statistics. We find six galaxies with $L_{\text{up}} = 10^{38.3-38.5} \text{ erg s}^{-1}$ and four with $L_{\text{up}} = 10^{38.9-39.1} \text{ erg s}^{-1}$, which appear to form a double-peak distribution. The truncation luminosity in individual galaxies is possibly correlated to the Hubble type T , as plotted in Figure 3. A tentative

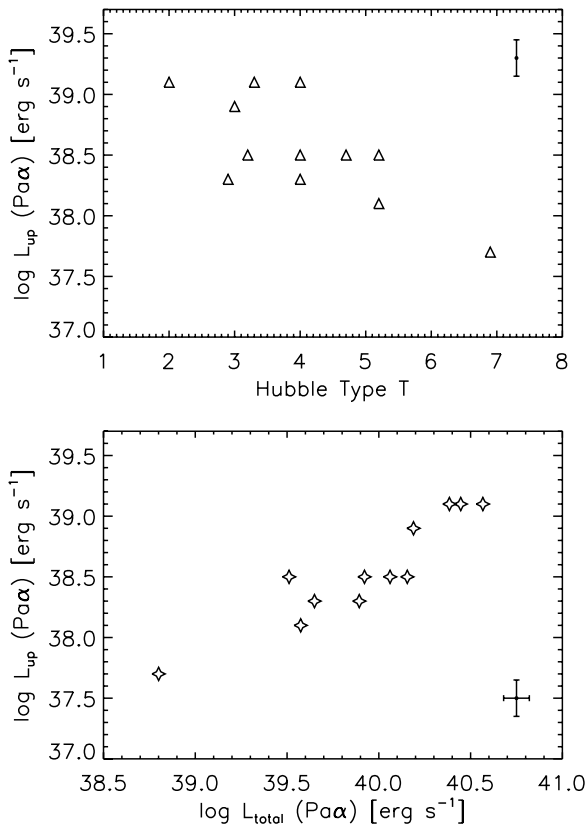


Figure 3. Dependence of the truncation luminosity of Pa α H II region luminosity function (L_{up}) on the morphological type and the total Pa α luminosity (L_{total}) of the host galaxy. The numerical Hubble type T of each galaxy is listed in Table 1, and L_{total} is the total luminosity of all the H II regions identified in each galaxy with $S/N \geq 3$ and listed in Table 2.

overall trend is observed, for which L_{up} is higher for galaxies with earlier morphological types, and the faintest data point, $L_{\text{up}} = 10^{37.7} \text{ erg s}^{-1}$, is contributed by the one with the latest type, the only Sd spiral NGC 3621. Our calculation shows that the Kendall rank correlation coefficient $\tau = -0.44$ with 5% probability that no correlation is present, but the correlation is tentative because it actually hinges on the single Sd galaxy. Meanwhile, we observe a tight correlation between L_{up} and the total luminosity of H II regions (L_{total}) in each galaxy (Figure 3), indicating an L_{total} versus T correlation which probably gives rise to the above relationship. However, the relatively small sample size prevents us from drawing a definitive conclusion from this data set. No dependence of L_{up} on the distance and inclination or rotation velocity of the sample galaxies is observed.

In Figure 2, we also find very few H II region candidates brighter than $L_{\text{Pa}\alpha, \text{br}}$, in agreement with previous *HST* results. On a galaxy-by-galaxy basis, their number is too small to infer any property (especially, NGC 278, NGC 3621, NGC 3810, and NGC 5248 do not have H II regions more luminous than $\sim 10^{38.5} \text{ erg s}^{-1}$ at all), so we will revisit this issue later, once we co-add all the LFs together.

3.2. Individual Galaxies: Luminosity–Size Relation

In order to characterize the spatial scale of an H II region with irregular morphology, we define an equivalent diameter (D) which matches the area of a circle $\pi D^2/4$ to that of the region. In general, larger H II regions tend to be brighter, leading to an empirical power-law scaling relationship between their luminosities and sizes, expressed as $L_{\text{Pa}\alpha} \propto D^\eta$. We show the luminosity–size correlation derived from our Pa α images

in Figure 4, where all $S/N \geq 3$ regions are fit to the power law $L_{\text{Pa}\alpha} \propto D^\eta$. We notice that all of the 12 galaxies show a value of η between 2 and 3, and most (10 out of 12) galaxies have $\eta \gtrsim 2.5$. The best-fit exponent η is independent of the distance of the galaxy. However, there exists the possibility that the tight luminosity–diameter scaling relations are not physical, but regulated by the surface brightness limits in the way we have defined objects. To test this, we run simulations by placing artificial sources onto the Pa α maps with Gaussian profiles, varying the luminosities and sizes of the sources, and running *HIIphot* using the same parameter setting as for the respective galaxies to see if they can be identified. Fainter and more extended H II regions are harder to detect, which forms an envelope below which the regions are not identified by the algorithm. We find that our observed luminosity–diameter relations are well above this limiting envelope for source identification, and are therefore physical rather than artificial. This envelope is shown for NGC 1097 as an example in Figure 4.

3.2.1. Source Blending Effects

The existence of tight luminosity–size scaling relations like those of Figure 4 begs for an explanation of the variations in the observed power-law index η , which generally has values between 2 and 3.

Scoville et al. (2001) argue that the luminosity is linked to D through $L_{\text{H}\alpha} \propto D^3$ if all the H II regions in a galaxy are resolved and no blending is present, while $L_{\text{H}\alpha} \propto D^2$ in the case of strong blending. Traditional ground-based H α imaging is limited to $\sim 1''$ – $2''$ by atmospheric seeing. At a distance of 20 Mpc, this angular scale translates to 100–200 pc, larger than the diameter of a typical H II region, although comparable to the size of the much rarer giant H II regions, like 30 Doradus in the LMC. One might expect that higher resolution data like those provided by the *HST* would break the blends with ease. However, Scoville et al. (2001) find an $L_{\text{H}\alpha} \propto D^2$ relation for M51, even when using the $0''.1$ resolution *HST* images (corresponding to a linear scale of 4.5 pc at the distance of M51). Those authors thus argue that source blending is still strong in M51, despite the high angular resolution of the data.

If source blending were the reason for flattening η toward a value of 2, it would be difficult to understand two observational facts. (1) Our Pa α observations yield larger η values than Scoville et al. (2001), suggesting that our data break the H II blends more effectively, despite their lower angular resolution ($0''.26$) relative to that of Scoville et al. (2001) in H α ($0''.1$). (2) When we divide the H α image obtained by Scoville et al. (2001) into nine annuli and run *HIIphot* on each of them, we find that the L – D relation becomes flatter as a function of the galactocentric distance (Figure 5). The best-fit value of η is 3.3 in the central regions and decreases monotonically to ~ 2 at $\gtrsim 4$ kpc. This result is counterintuitive because the central kpc region of M51 is significantly more crowded than its outer regions, as promptly seen in the H α image.

In fact, the interpretation of the $L \propto D^2$ relation as an effect of blending is questionable. If individual regions (denoted i) have luminosity–diameter relations $L_i \propto D_i^3$, the luminosity–diameter relation of the blended regions will also show an overall $L \propto D^3$ trend. In reality, as both L_i and D_i will cover a range of values, the blended luminosity–size relation is likely to have a more complicated functional form. In order to derive the dependence of η on region blending, we carry out a simple simulation by generating 10^6 H II regions whose luminosities L_i are randomly distributed between 10^{35}

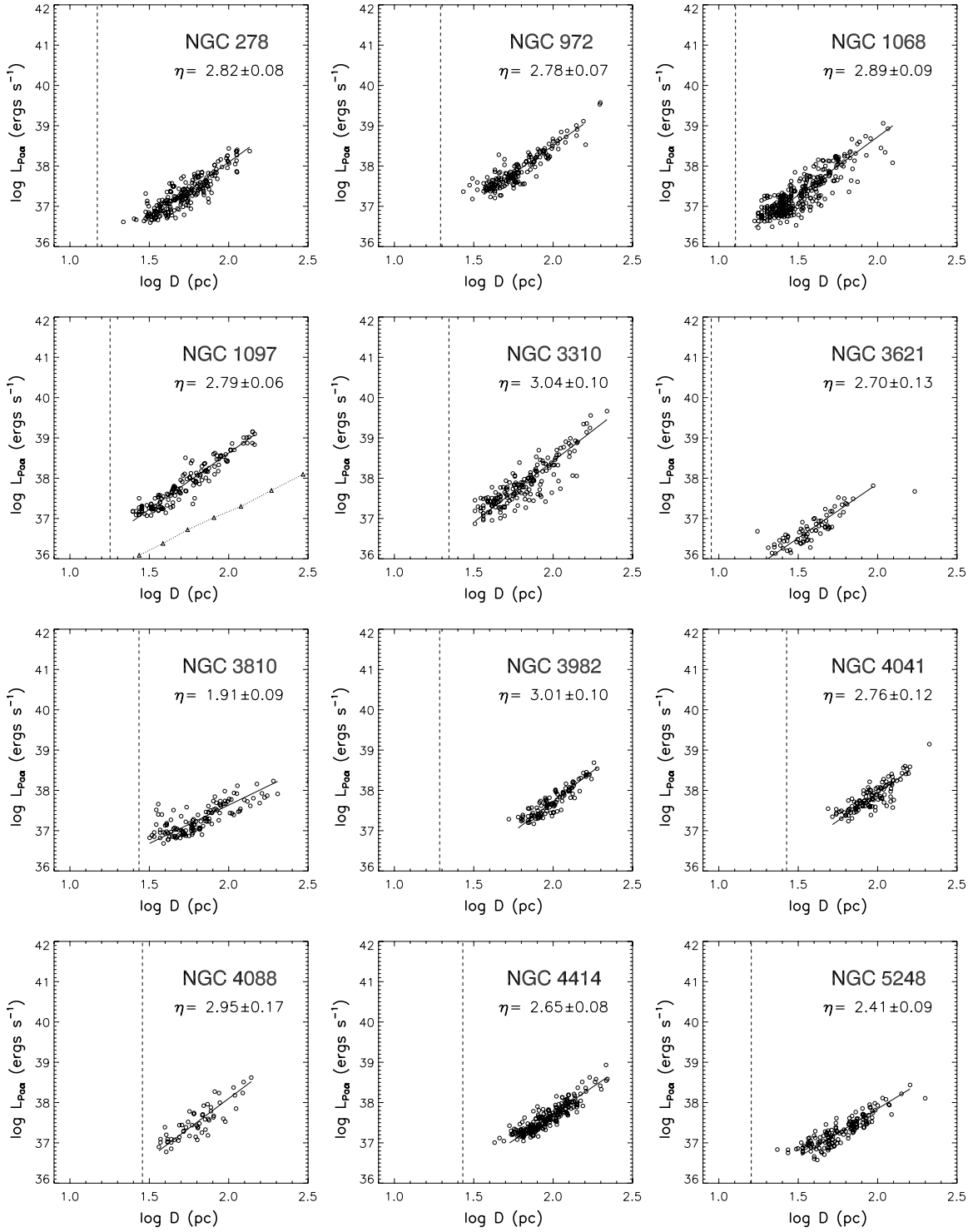


Figure 4. Observed Pa α luminosities of the H II regions in the individual galaxies as a function of equivalent diameter (D). The power-law fits are performed on regions with $S/N \geq 3$. The linear resolution listed in Table 1 is shown by a vertical dashed line for each galaxy. In the panel for NGC 1097, the limiting envelope for source identification is shown by triangles connected with dotted lines (see the text).

and $10^{38} \text{ erg s}^{-1}$ and following $dN(L) \propto L^{-2} dL$. We divide the whole sample into subsamples containing regions each resulting from n blends. After calculating L and D using the same formalism as Scoville et al. (2001),

$$L = \sum_{i=1}^n L_i \propto \sum_{i=1}^n D_i^3, \quad \text{and} \quad D = \left(\sum_{i=1}^n D_i^2 \right)^{1/2}, \quad (2)$$

we fit a power law to the simulated data and find that the exponent η actually increases with the number n of blended regions. Specifically, η is found to range from 3.00 to 3.81 for n that progressively increases from 1 to 10. Therefore, source blending steepens the L - D relation and leads to $\eta > 3$, instead of reducing η to a value ~ 2 as suggested by Scoville et al. (2001).

This simulation also provides an explanation for the larger values of η we measure in our individual galaxies relative to that measured by Scoville et al. (2001) in M51: blending

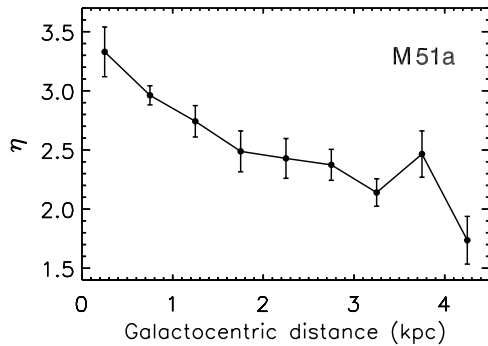


Figure 5. Best-fit exponent of the luminosity–size relation ($L \propto D^\eta$) of the M51 H II regions, plotted as a function of the galactocentric distance on the star-forming disk. This analysis is performed on the same *HST* H α image as that used by Scoville et al. (2001).

actually steepens the luminosity–size relation. Furthermore, the radial profile of η in M51, with decreasing values for increasing galactocentric distance, can be understood as a natural result of decreasing source blending effects. However, we still have not explained why $\eta \approx 2$ in many observational studies.

3.2.2. Other Effects

In this subsection, we explore other effects that may drive the observed values of η : (1) differences between the H α and Pa α analyses, (2) the H II region identification and measurement algorithm, (3) the presence of noise in the data, or (4) a faint-end cutoff in the H II LF. These may all play a role in shaping the L – D relation and could counteract the steepening effect of source blending. To disentangle the interplay among these factors, we have carried out a series of experiments.

In order to test the consistency of the H α and Pa α analyses, we compare the LFs and the luminosity–size relations derived from M51a for which images in both lines are available. The central $R \leq 2.75$ kpc region of M51a is mapped both in H α and Pa α by *HST* (Scoville et al. 2001). Running *HIIphot* over this region results in an H α LF with $\alpha_{\text{H}\alpha} = -1.12 \pm 0.05$ and a Pa α LF with $\alpha_{\text{Pa}\alpha} = -1.09 \pm 0.05$, which are fully consistent with each other. Using *HIIphot* again, this time on the entire FOVs of the *HST* H α and Pa α images of M51a, to reconstruct the L – D relation, we find best-fit power indices $\eta_{\text{H}\alpha} = 2.47 \pm 0.04$ for the entire covered disk with a diameter of ~ 8 kpc and $\eta_{\text{Pa}\alpha} = 2.37 \pm 0.07$, consistent with each other within 1σ . We thus recover a steeper slope than either Scoville et al. (2001), who derive $\eta_{\text{H}\alpha} = 2.1$, or Gutiérrez et al. (2011), who find $\eta_{\text{H}\alpha} = 1.924$. Our relation also shows significantly more scatter than that found by both Scoville et al. (2001) and Gutiérrez et al. (2011) (both use similar algorithms to identify and recover H II regions), but similar to the scatter recovered by Lee et al. (2011), who use *HIIphot* (but derive $\eta_{\text{H}\alpha} = 2.16 \pm 0.02$). Clearly, the employed algorithm for identifying and measuring the H II regions has an effect on the level of scatter recovered in the L – D relation.

When limiting our analysis to the central 5 kpc diameter of M51 (the largest area available for our 12 program galaxies), we find an H α L – D relation with $\eta_{\text{H}\alpha} = 2.62 \pm 0.07$. This value is consistent with our Pa α results that yield an average exponent $\langle \eta \rangle = 2.70 \pm 0.27$, implying consistency between our Pa α analysis and the H α analysis of Scoville et al. (2001). We conclude that although our Pa α data are shallower than the M51 H α by a factor of about 10 (Kennicutt et al. 2007), the derived statistical properties of H II regions are consistent with each other. The use of different emission lines is ruled out as a driver for $\eta \sim 2$.

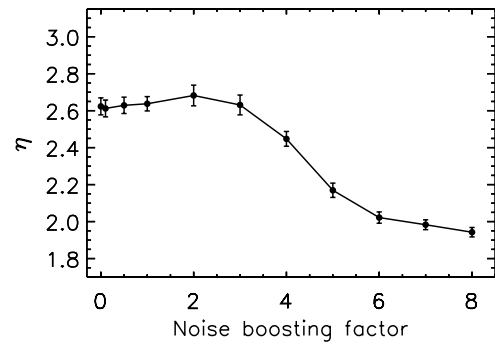


Figure 6. Power-law index of the H II region luminosity–size relation as a function of the noise level, tested using simulated images. The noise is extracted from the Pa α image of NGC 1097 and multiplied by the noise boosting factor before being added to the simulated H II regions.

We now proceed to quantify the effect of using a specific algorithm for identifying and measuring H II regions (in our case: *HIIphot*) and how this plays out for varying noise levels in the data. We carry out the experiment detailed below by running *HIIphot* on a set of simulated Pa α images. We first create a noise map with the same size and pixel scale as our Pa α images by mosaicking the blank sky regions of the NGC 1097 Pa α map repeatedly. After that, we generate a sample of 800 H II regions whose luminosities are randomly drawn from the distribution $dN(L)/d \ln L \propto L^{-1}$ in the range 10^{35-39} erg s $^{-1}$. The upper limit is anchored to the maximum H II region luminosity in NGC 1097 but is roughly applicable to other galaxies as well. The luminosity range of the observed H II regions covers about two orders of magnitude (Figure 4), but we adopt a much broader range so that the faint end extends well below the detection limit of the H II regions. This is to avoid that the low-luminosity end may play a role in the results. Each region is represented by a two-dimensional Gaussian profile whose integrated flux is the assigned luminosity and whose FWHM satisfies $L \propto \text{FWHM}^3$ (the coefficient is adjusted to roughly match the observed size range). When we assign a random position to each region, we create a noiseless simulated image. In order to examine the variation of the L – D relation at different noise levels, we multiply the noise map by a “noise boosting factor” which varies from 0, 0.1, 0.5 and 1, 2, 3, ... to 8 progressively before adding it to the noiseless image to create a set of final simulated images. Using a parameter setting identical to what was applied to NGC 1097, we apply *HIIphot* on the simulated images, and the measured values of η are plotted in Figure 6.

From this figure, we conclude that $\eta < 3$ results from a combination of the *HIIphot* algorithm and the presence of noise in the data. Even when the noise boosting factor is 0 (noiseless data), we do not retrieve the input $L \propto D^3$, but find a significantly flatter slope ($\eta = 2.62 \pm 0.05$), close to the average value in our sample, $\langle \eta \rangle = 2.70 \pm 0.27$. Thus, the H II region retrieval algorithm has a measurable effect on the recovered η . As the noise increases, η starts decreasing, reaching a value of ~ 2 , when the noise is increased by a factor of six or more. From these simulations, we conclude that the value of η decreases and reaches an asymptote of ~ 2 as the noise in the image increases, or, analogously, as the S/N of the H II regions decreases.

If the minimum H II region that can be detected in a galaxy is ionized by a cluster containing a single O7.5 star, the corresponding Pa α luminosity is $10^{36.2-36.3}$ erg s $^{-1}$. A faint-end luminosity cutoff has also the property of flattening the L – D relation, almost irrespective of the level of blending.

Running Monte Carlo simulations, we recover smaller values of η corresponding to $\Delta\eta \sim 0.4$ – 0.5 , for the cutoff above, even when the data are noiseless. Higher cutoff luminosities yield stronger flattening for η .

In summary, the exponent of the L – D relation, η , is subject to variations that are unrelated to intrinsic physical changes in the properties of the H II regions, but are most likely induced by observational biases, as already noted by Scoville et al. (2001). However, not all observational biases work in the direction of yielding smaller-than-true values of η . For instance, our simulations indicate that source blending steepens the L – D relation to values $\eta > 3$. Conversely, η flattens to values < 3 and approaches $\eta \sim 2$ in the presence of both decreasing S/N data and a faint-end cutoff to the LF of H II regions. The use of the algorithm *HIIphot* to identify and measure the H II regions also has some role in recovering values of $\eta < 3$. Instead, the use of either H α or Pa α images to perform the analysis has negligible impact on the recovered values of η .

In light of the above, we can interpret both $\eta \sim 2$ as measured by Scoville et al. (2001) in M51 and the radial profile of η in the same galaxy (Figure 5). In the center of the galaxy, significant blending yields $\eta > 3$, but, as the galactocentric distance increases, both source blending and the overall S/N in the image decrease, thus reducing η . The presence of an intrinsic faint-end cutoff in the H II regions of M51 would further help reduce η . In spite of these artifacts, the intrinsic radial dependence of the H II region properties and their environment, if it exists, could also play a role in shaping the radial profile of η .

3.3. Size and Electron Density Distribution

The H II region size distributions for the sample galaxies are plotted in Figure 7 for regions with S/N ≥ 5 , and the large-size end is fitted with a power law formulated similarly to that of the Pa α line luminosity:

$$dN(D)/d \ln D \propto D^\beta. \quad (3)$$

For isolated sources, $L \propto D^3$ is expected, and the power index in the above equation will take the value $\beta = -3$ (Oey et al. 2003; note that β is linked to the parameter b therein, which satisfies $N(D)dD \propto D^{-b}dD$, through $\beta = 1 - b$). In case of $L \propto D^2$ as often seen in observations, $\beta = -2$ instead. The H II region size distribution in most of our galaxies is consistent with a slope $\beta \sim -3$ (Figure 7), but variations from galaxy to galaxy are significantly larger than for the LF power index α . For instance, the size distributions of NGC 1097 and NGC 3982 do not appear to follow a power-law behavior at all. Averaging over the sample without including these two objects, the mean value of the exponent is $\langle \beta \rangle = -3.42$, and the median $\tilde{\beta} = -3.10$, with an uncertainty of ± 1.04 , consistent with $\beta = -3.12 \pm 0.90$ found by Oey et al. (2003). The largest spatial scale beyond which the power-law scaling truncates (presumably because of small number statistics), D_{up} varies from ~ 90 pc to ~ 200 pc.

The remarkably larger uncertainty on β is partly caused by the fact that the H II region sizes are much more weakly correlated with S/N than the luminosities. For this reason, applying an S/N threshold only leads to a sharp cutoff at the faint end of the LF, but changes the size distribution *globally* or even artificially destructs its power-law scaling. This shortcoming can only be fixed with future deeper data. On the other hand, one should be aware that the above simple calculations toward an expected behavior of the size distribution are based on an implicit

assumption that the three relations (the H II LF, luminosity–size correlation, and size distribution) are all dominantly driven by the same objects. However, as can be seen in Figures 2 and 4, the values of α and η are basically determined by the full sample of all detected H II regions (the power-law scaling roughly extends to completeness limits in most galaxies; see Figure 2), but β is only determined by a fraction (often $< 50\%$) of the largest regions (Figure 7). We will see in the co-added analysis of the H II region size distribution (Section 4 and Figure 11) that the largest H II regions (some possibly blends) form a power law with $\beta = -3.89 \pm 0.48$.

Assuming an intrinsic line ratio $H\alpha/\text{Pa}\alpha = 8$ and an electron temperature $T_e = 7500$ K (see Section 2.3), we express the average electron density of an H II region as (Scoville et al. 2001, Equation (5))

$$\frac{n_e}{\text{cm}^{-3}} = 34 \left(\frac{L_{\text{Pa}\alpha, \text{intr}}}{10^{36} \text{ erg s}^{-1}} \right)^{1/2} \left(\frac{D}{10 \text{ pc}} \right)^{-3/2}, \quad (4)$$

where the intrinsic Pa α luminosity is employed. As shown in Figure 8, the mean number density of electrons in each galaxy spans over the range $\langle n_e \rangle = 8$ – 28 cm^{-3} in our sample. With a typical dust extinction in H II regions of $A_V = 2.2$ mag (Calzetti et al. 2007), we expect the intrinsic n_e to be larger by a factor of ~ 1.2 assuming the Milky Way extinction curve (Cardelli et al. 1989). The intrinsic mean electron density of the H II regions in our sample galaxies is therefore in the range 10 – 33 cm^{-3} , similar to the extinction-corrected $n_e = 15$ – 60 cm^{-3} found in M51 (Scoville et al. 2001). However, we emphasize that the electron densities derived here are rms densities averaged over the entire H II regions. The actual local densities are often significantly higher than these values because H II regions tend to be highly clumpy and filamentary with filling factors of gas of order 1% or less.

4. CO-ADDED ANALYSIS

In H II region analyses of individual galaxies, one of the primary disadvantages of using Pa α compared to H α data is the smaller number statistics, driven by the small FOV of *HST*/NICMOS. Aiming for a better statistical significance, we co-add the H II region catalogs of our galaxies and analyze the combined sample. However, we should caution that the co-added analyses implicitly employ a strong assumption that a single universal LF exists in all involved galaxies. For our galaxy sample with a median value for the LF power index $\tilde{\alpha} = -0.98 \pm 0.19$, a Kolmogorov–Smirnov test implies a probability of 85.1% that the 12 measured α values are drawn from a Gaussian probability density function with $\mu = -0.98$ and $\sigma = 0.19$.

The co-added LF is shown in Figure 9. The completeness limit of this combined H II catalog, $10^{37.78} \text{ erg s}^{-1}$, is set by NGC 4041, which has the shallowest image among the involved galaxies (in terms of H II region luminosity). Above this limit, the locus of the distribution can be fitted by a single power law with an exponent of $\alpha = -1.07 \pm 0.07$ spanning over the range $10^{37.8-39.4} \text{ erg s}^{-1}$ (Figure 9(a)). Although we do not observe convincing evidence for the presence of a broken power law in our data, a potential break is seen at $L_{\text{Pa}\alpha, \text{br}} \sim 10^{38.5} \text{ erg s}^{-1}$, corresponding to an observed (i.e., not corrected for extinction) H α luminosity of $L_{\text{H}\alpha, \text{br}} \sim 10^{38.6-38.8} \text{ erg s}^{-1}$, consistent with the transition point $L_{\text{H}\alpha, \text{br}} = 10^{38.6 \pm 0.1}$ determined by Bradley et al. (2006). In Figure 9(b), we adopt this break luminosity

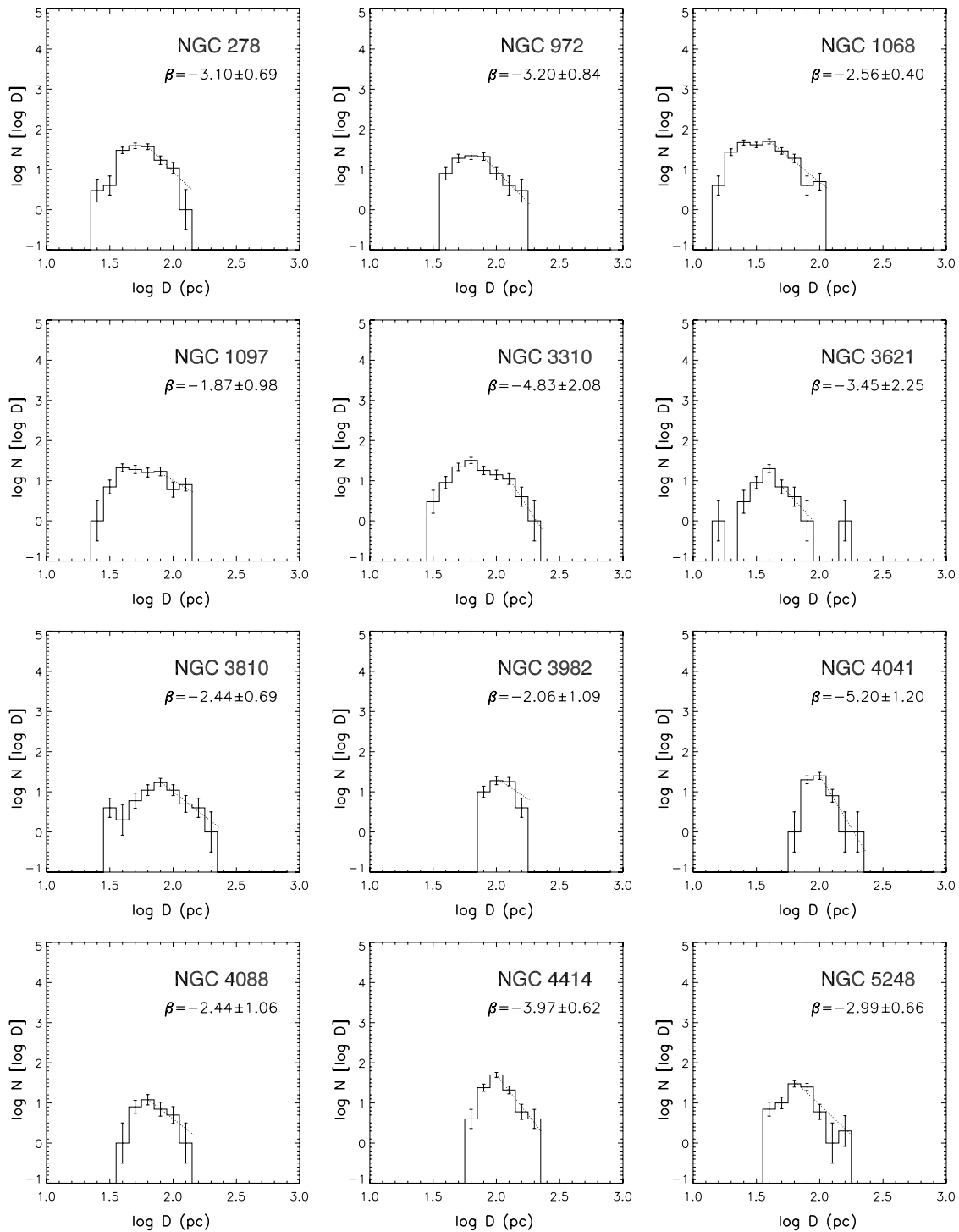


Figure 7. H II region size distribution of the 12 sample galaxies. The power-law fits are performed in a manner similar to Figure 2. Only regions with $S/N \geq 5$ are shown of which the large-size end is fitted.

and fit the regimes above and below this point, separately. We emphasize that this translation depends on our assumed constant extinction value. That is, if the extinction is lower than $A_V = 2.2$ (see Section 2.3), then the $\text{Pa}\alpha$ luminosity break will be considerably lower than the $\text{H}\alpha$ one for the same population.

We find that the best-fit power index is $\alpha_{\text{br}+} = -1.14 \pm 0.43$ for the regions beyond $L_{\text{Pa}\alpha, \text{br}}$, and $\alpha_{\text{br}-} = -0.90 \pm 0.11$ for the fainter ones. These two indices are both approximately -1 when the uncertainties are taken into account, steeper (especially $\alpha_{\text{br}-}$) than those found by Bradley et al. (2006), $\alpha_{\text{br}+} = -0.86$

and $\alpha_{\text{br}-} = -0.36$. The one order of magnitude lower angular resolution of their data (seeing up to $3''.7$ or ~ 300 pc) is likely responsible for this difference. Another difference relative to previous results is that, despite that the breaks in our co-added LF and that of Bradley et al. (2006) are both located at roughly the same luminosity, our two power laws are not only “broken” but also “discontinuous.” This discontinuity is sometimes also vaguely seen in the $\text{H}\alpha$ analyses of some individual galaxies (e.g., NGC 5068, NGC 6300, etc.; see Helmboldt et al. 2005).

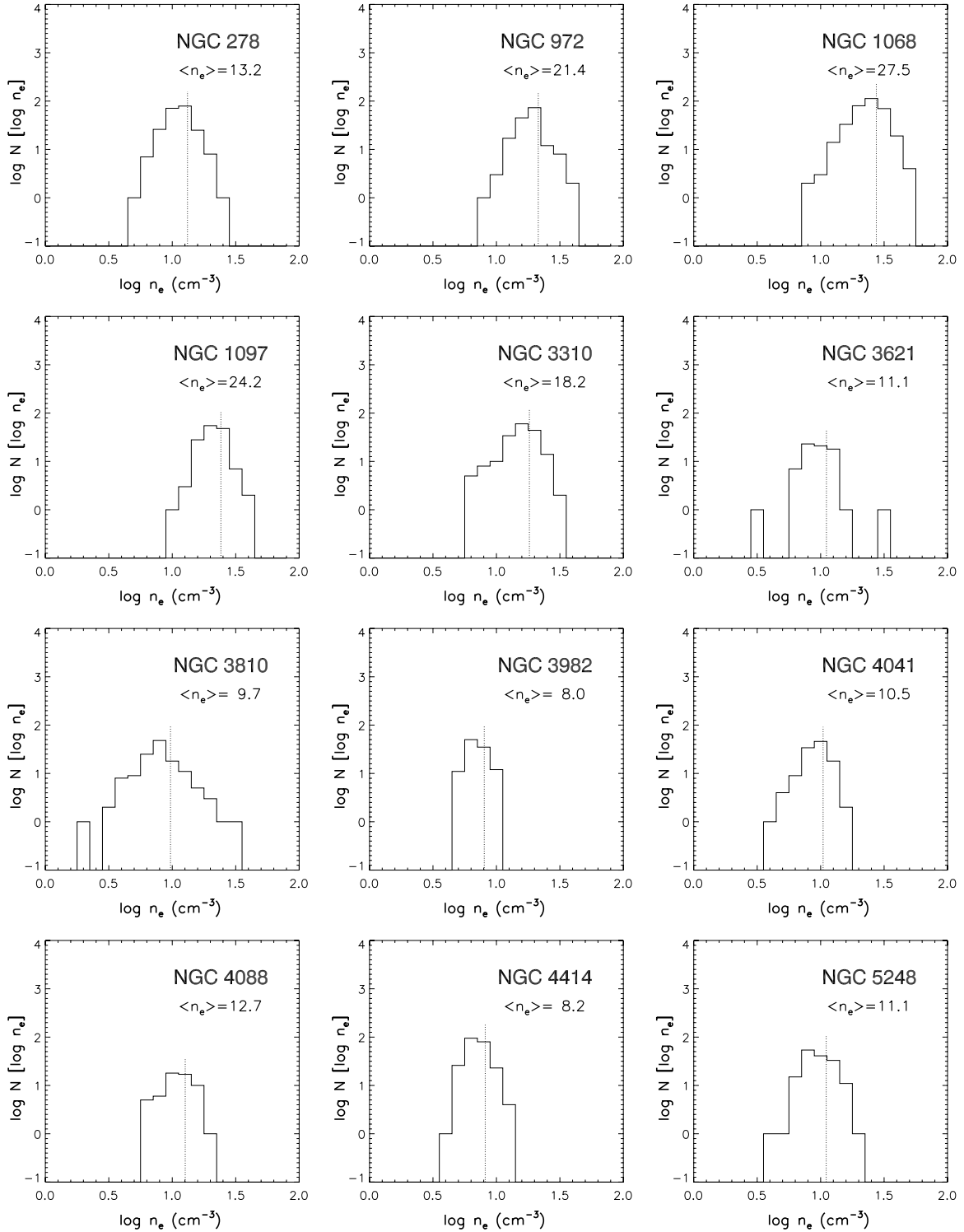


Figure 8. Distribution of the mean electron density (n_e) of H II regions in individual galaxies. Note that the gas in H II regions is clumpy and filamentary with local densities often much higher than the rms densities measured here. The electron density averaged over each galaxy ($\langle n_e \rangle$) is depicted by vertical dotted lines.

It would be intriguing if the parameter(s) that may differentiate the H II regions that populate the nearly parallel double power laws could be tracked down. As an exploratory experiment, we set a cutoff scale D_{cut} , construct the LF excluding regions with $D > D_{\text{cut}}$, and observe its corresponding variation by successively changing D_{cut} . When D_{cut} decreases from the largest observed value of D , we see that the regions of the LF locus that gets more drastically depopulated are those at the bright end, until nearly all regions disappear when $D_{\text{cut}} =$

100 pc. During this process, the power scaling on the faint side of $L_{\text{Pa}\alpha, \text{br}}$ remains almost unchanged, both in shape and index. When D_{cut} decreases further, the fainter locus starts to lose objects in a significant manner that starts changing the power-law distribution. We show in Figure 10 the co-added LF populated by H II regions smaller than 100 pc in diameter. We therefore consider ~ 100 pc characteristic scale that differentiates the two H II region populations which construct their respective power scalings on either side of the break luminosity,

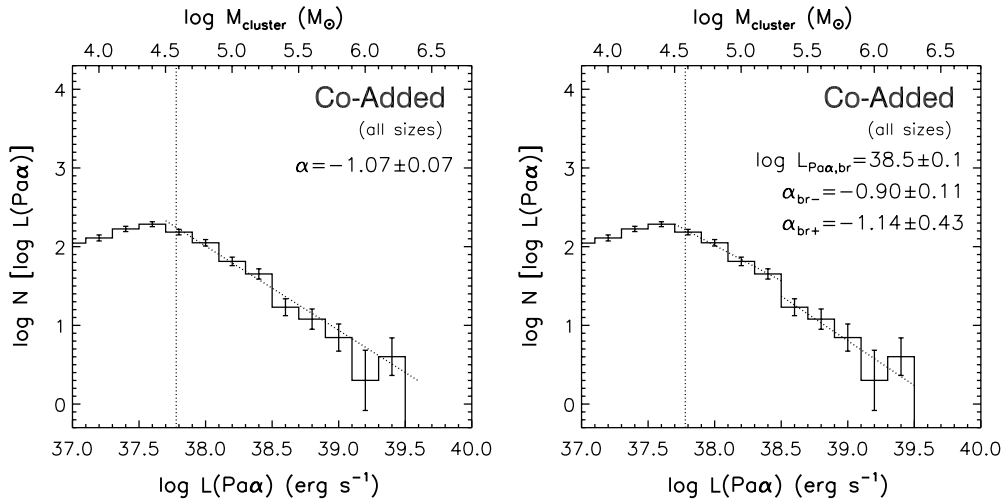


Figure 9. Co-added $\text{Pa}\alpha$ luminosity function of H II regions for all objects above the completeness limit as denoted by the vertical dotted lines, fitted by a single (left) or double (right) power law. The mass of the star cluster as a function of $L_{\text{Pa}\alpha}$ is calculated using the stellar population synthesis model *Starburst99* (Leitherer et al. 1999) assuming solar metallicity, an age of 4 Myr, and the IMF given by Kroupa (2001) in the range 0.1–120 M_{\odot} .

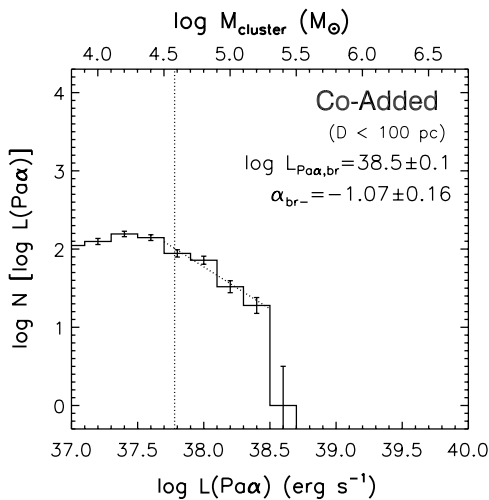


Figure 10. Co-added $\text{Pa}\alpha$ luminosity function of H II regions for all objects as significant as in Figure 9, but with a diameter smaller than 100 pc only. The mass of the star cluster is calculated the same way as in Figure 9.

if the break is real. Excluding regions that have sizes >100 pc, the power index of the co-added LF for sub-break regions becomes slightly steeper ($\alpha_{\text{Pa}\alpha, \text{br}-} = -1.07 \pm 0.16$) than otherwise ($\alpha_{\text{Pa}\alpha, \text{br}-} = -0.90 \pm 0.11$) and slightly more consistent with the mean slope for individual galaxies, $\langle \alpha \rangle = -1.05$. However, the discrepancy of the data from a single power law is at the level of $\sim 1\sigma$ uncertainty of a typical data point and is therefore insignificant. Although a fraction of the regions larger than ~ 100 pc are probably blends, the existence of large H II regions is unquestionable and we cannot arbitrarily attribute the break to source blending effects.

We have mentioned in Section 3.1 that H II regions more luminous than $\sim 10^{38.5} \text{ erg s}^{-1}$ are not seen in four galaxies. Although excluding these objects from the co-added analysis will impose a further criterion that may bias our sample, including them for testing the existence of the break raises concerns, as well. In fact, we find that removing these galaxies from the sample will only flatten the sub-break power index slightly ($\alpha_{\text{br}-} = -0.82 \pm 0.11$), and the (in)significance of the break and discontinuity remains unchanged, because the

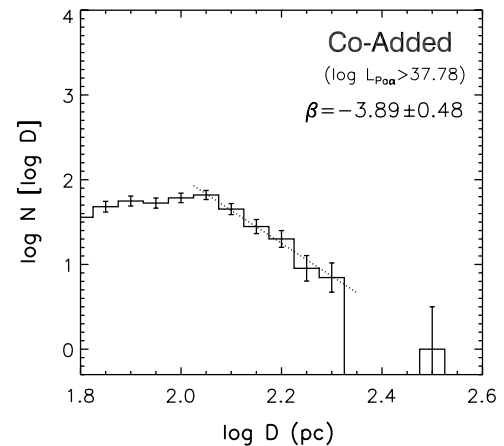


Figure 11. Co-added size distribution of H II regions for all objects above the completeness limit. A well-shaped power law exists over the range 110–220 pc although some of the largest regions are possibly blends.

majority of the H II regions contributed by these four galaxies are fainter than the completeness limit ($10^{37.78} \text{ erg s}^{-1}$).

We show in Figure 11 the co-added size distribution of the H II regions in our sample galaxies, excluding all the regions fainter than the completeness limit, $10^{37.78} \text{ erg s}^{-1}$, set by the shallowest image, i.e., that of NGC 4041. A well-shaped power law exists over the range 110–220 pc, with a slope $\beta = -3.89 \pm 0.48$, consistent with the individual-galaxy analysis presented above. Although some regions at the large-size end are possibly blended regions or complexes, this locus may still reflect the intrinsic clustering properties of compact H II regions to some extent.

5. DISCUSSION

The LFs of H II regions ionized by OB star associations are physically related to the mass function of compact star clusters and the mass spectrum of giant molecular clouds in galaxies because young stars are formed predominantly in Giant Molecular Clouds (GMCs) and ionize the surrounding interstellar medium to produce gaseous nebulae.

For a constant and universal stellar IMF, the number of ionizing photons emitted by the short-lived, massive stars is proportional to the total mass of such stars, which in turn is

proportional to the total mass of the star cluster, if the latter is sufficiently young that no supernovae have exploded yet (age $< 3\text{--}4$ Myr). In this case, the compact cluster mass function should follow the same power law as the H II region LF. Recent *HST* studies find that $dN/dM_\star \propto M_\star^{-2}$ (Chandar et al. 2010, 2011), for clusters around 10^8 Myr, which is a power law similar to that found for the H II region LF. This result would argue for cluster destruction/dissolution mechanisms that are mass-independent (Chandar et al. 2010) in order to preserve the constant power-law shape between the younger H II regions' LF and the older clusters' mass function.

The relation between H II LFs and GMC mass spectra (expressed as $dN/dM_{\text{GMC}} \propto M_{\text{GMC}}^\zeta$) is more sophisticated because the efficiency of star formation is likely a function of the local environment. Empirically, the SFR is linked to the density of cold gas through a single power-law scaling relation (the Schmidt–Kennicutt law, abbreviated as the S-K law; see Kennicutt 1998 and references therein) which takes the form $\Sigma_{\text{SFR}} \propto \Sigma_{\text{gas}}^\gamma$. Ideally, the extragalactic S-K law should be studied for each individual cluster and its parent GMC, and the *molecular* S-K law should be expressed as $\text{SFR} \propto L_{\text{Pa}\alpha} \propto M_{\text{GMC}}^{\gamma_{\text{H2}}}$, if observed in Pa α . For the power-law Pa α H II LF that we find in this work, the GMC mass spectra would have an exponent of $\zeta = -(1 + \gamma_{\text{H2}})$. Although γ_{H2} depends on the sampling scale as shown by Liu et al. (2011) and Calzetti et al. (2012), for the case of $\gamma_{\text{H2}} = 1.4$ as found in M51 at 750 pc resolution by Kennicutt et al. (2007), this relation predicts $\zeta = -2.4$, in rough agreement with the observed exponent for the GMC mass function. Recent observations have found the value of ζ to be $-(1.5\text{--}2.1)$ in the Galaxy, -1.7 in LMC, and -2.9 in M33, therefore varying significantly across the Local Group, but a truncation at the maximum mass of $10^{6.5} M_\odot$ exists (Rosolowsky 2005). The above calculation, assuming that all molecular clouds can be well detected, well separated, are all forming stars, and each can be clearly identified as the parent cloud of a star cluster, is far from being realistic and is presented as heuristic discussion only.

For an H II region with a Pa α luminosity of $10^{38.5} \text{ erg s}^{-1}$, the corresponding star cluster has a mass of $(2\text{--}3) \times 10^5 M_\odot$ for ages 4–5 Myr, if calculated using the stellar population synthesis model *Starburst99* (Leitherer et al. 1999), assuming the IMF given by Kroupa (2001) in the range $0.1\text{--}120 M_\odot$ and solar metallicity. If such a star cluster is formed in a GMC with the maximum mass, $10^{6.5} M_\odot$, the star formation efficiency will be 10% or less, consistent with Leroy et al. (2008) and Bigiel et al. (2008) (note that the bright end of the S-K law is considered here, where the influence of preserving the diffuse stellar/dust emission mostly due to old stellar populations is insignificant; see Liu et al. 2011 and Calzetti et al. 2012 for detailed discussion on the influence of the local diffuse emission on deriving spatially resolved star formation properties of nearby galaxies).

6. SUMMARY

We have analyzed the *HST*/NICMOS Pa α images of 12 nearby galaxies and studied the LF and size distribution of the H II regions both in individual galaxies and cumulatively. Our analysis is the first using a sample of galaxies, rather than individual objects, with *HST* data. The use of high angular resolution data offers an advantage over ground-based investigations (mainly in H α) because it avoids severe blending (spatially or in projection) of H II regions.

The near-IR Pa α hydrogen line is virtually an extinction-free trace of newly formed massive stars and star associations. The resolution of our Pa α data ($0''.26$) translates to 9–29 pc at the distances of our sample galaxies (7–23 Mpc), satisfying the requirement for resolution to reliably measure the properties of H II regions (< 40 pc; Pleuss et al. 2000). Employing the *Hilphot* IDL software for the identification and photometry of H II regions, we show that the luminosity–diameter relation of our cataloged H II regions is characterized by $L \propto D^\eta$ with $2.5 < \eta < 3$ for 10 galaxies.

In the individual-galaxy analysis, we do not confirm the double power-law LF found by typical ground-based works. Instead, all 12 galaxies exhibit Pa α LFs that follow a single power law $dN(L_{\text{Pa}\alpha})/d \ln L_{\text{Pa}\alpha} \propto L_{\text{Pa}\alpha}^\alpha$ with $\alpha \simeq -1$, consistent with the investigations of both Galactic radio H II regions and the *HST* studies of M51 and M101. For the whole sample, we find a median value for the LF power index $\tilde{\alpha} = -0.98 \pm 0.19$. The size distribution of H II regions, fitted to $dN(D)/d \ln D \propto D^\beta$, shows significantly larger variation of the exponent than the LF, with two galaxies showing no apparent power scaling at all. Excluding these two objects, we find the median value $\tilde{\beta} = -3.10 \pm 1.04$, consistent with $\beta = -3.12 \pm 0.90$ found by Oey et al. (2003).

The co-added LF can be fitted by a single power law with an exponent of $\alpha = -1.07 \pm 0.07$ over a luminosity range of 1.6 orders of magnitude, but a possible break is seen at $L_{\text{Pa}\alpha, \text{br}} \sim 10^{38.5} \text{ erg s}^{-1}$, roughly corresponding to the transition luminosity determined by ground-based H α observations. When fitted with a disconnected double power law, the co-added LF has a best-fit power index $\alpha_{\text{br}+} = -1.14 \pm 0.43$ for the brighter regions and $\alpha_{\text{br}-} = -0.90 \pm 0.11$ for the fainter ones. The power scaling of the brighter regions is mainly populated by regions with diameters larger than 100 pc, which may be contaminated with blends. Excluding these objects results in a single power law with $\alpha = -1.07 \pm 0.16$. We conclude that the LFs, both individual and co-added, are consistent with a single power law with $\alpha = -1$, as expected from the mass function of star clusters in nearby galaxies. The co-added size distribution shapes a power law with a slope $\beta = -3.89 \pm 0.48$ over the range 110–220 pc. This scaling, in spite of its consistency with the individual-galaxy analysis, is probably contaminated by blended regions or complexes.

We thank the anonymous referee for a careful reading of the manuscript. Partial support for program GO-11080 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. We also acknowledge the usage of the HyperLeda database (<http://leda.univ-lyon1.fr>).

Facility: *HST* (NICMOS)

REFERENCES

- Azimlu, M., Marciniak, R., & Barnby, P. 2011, *AJ*, **142**, 139
- Beckman, J. E., Rozas, M., Zurita, A., Watson, R. A., & Knapen, J. H. 2000, *AJ*, **119**, 2728
- Bigiel, F., Leroy, A., Walter, F., et al. 2008, *AJ*, **136**, 2846
- Bradley, T. R., Knapen, J. H., Beckman, J. E., & Folkes, S. L. 2006, *A&A*, **459**, L13

- Bresolin, F., Garnett, D. R., & Kennicutt, R. C., Jr. 2004, *ApJ*, **615**, 228
- Caldwell, N., Kennicutt, R., Phillips, A. C., & Schommer, R. A. 1991, *ApJ*, **370**, 526
- Calzetti, D. 2007, *NCimB*, **122**, 971
- Calzetti, D., Kennicutt, R. C., Engelbracht, C. W., et al. 2007, *ApJ*, **666**, 870
- Calzetti, D., Liu, G., & Koda, J. 2012, *ApJ*, **752**, 98
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, **345**, 245
- Chandar, R., Whitmore, B. C., Calzetti, D., et al. 2011, *ApJ*, **727**, 88
- Chandar, R., Whitmore, B. C., Kim, H., et al. 2010, *ApJ*, **719**, 966
- de Vaucouleurs, G., de Vaucouleurs, A., Corwin, H. G., et al. 1991, Third Reference Catalogue of Bright Galaxies, Vols. 1–3 (Berlin: Springer)
- Gutiérrez, L., Beckman, J. E., & Buenrosto, V. 2011, *AJ*, **141**, 113
- Helmholtz, J. F., Walterbos, R. A. M., Bothun, G. D., & O’Neil, K. 2005, *ApJ*, **630**, 824
- Helmholtz, J. F., Walterbos, R. A. M., Bothun, G. D., O’Neil, K., & Oey, M. S. 2009, *MNRAS*, **393**, 478
- Ho, L. C., Filippenko, A. V., & Sargent, W. L. W. 1997, *ApJS*, **112**, 315
- Hodge, P., Lee, M. G., & Kennicutt, R. C., Jr. 1989, *PASP*, **101**, 32
- Hodge, P. W., Balsley, J., Wyder, T. K., & Skelton, B. P. 1999, *PASP*, **111**, 685
- Hsieh, P.-Y., Matsushita, S., Liu, G., et al. 2011, *ApJ*, **736**, 129
- Kennicutt, R. 1992, in *Star Formation in Stellar Systems* (Cambridge: Cambridge Univ. Press), 191
- Kennicutt, R. C., Jr. 1998, *ARA&A*, **36**, 189
- Kennicutt, R. C., Jr., Calzetti, D., Walter, F., et al. 2007, *ApJ*, **671**, 333
- Kennicutt, R. C., Jr., Edgar, B. K., & Hodge, P. W. 1989, *ApJ*, **337**, 761
- Knapen, J. H. 1998, *MNRAS*, **297**, 255
- Kroupa, P. 2001, *MNRAS*, **322**, 231
- Lee, J. H., Hwang, N., & Lee, M. G. 2011, *ApJ*, **735**, 75
- Leitherer, C., & Heckman, T. M. 1995, *ApJS*, **96**, 9
- Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, *ApJS*, **123**, 3
- Leroy, A. K., Walter, F., Brinks, E., et al. 2008, *AJ*, **136**, 2782
- Li, Y., Bresolin, F., & Kennicutt, R. C., Jr. 2013, *ApJ*, **766**, 17
- Liu, G., Koda, J., Calzetti, D., Fukuhara, M., & Momose, R. 2011, *ApJ*, **735**, 63
- McKee, C. F., & Williams, J. P. 1997, *ApJ*, **476**, 144
- Ngeow, C., & Kanbur, S. M. 2006, *ApJL*, **642**, L29
- Oey, M. S., & Clarke, C. J. 1998, *AJ*, **115**, 1543
- Oey, M. S., Parker, J. S., Mikles, V. J., & Zhang, X. 2003, *AJ*, **126**, 2317
- Osterbrock, D. E., & Ferland, G. J. 2006, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei* (2nd ed.; Sausalito, CA: Univ. Science Books)
- Paturel, G., Petit, C., Prugniel, P., et al. 2003, *A&A*, **412**, 45
- Petit, H. 1998, *A&AS*, **131**, 317
- Pleuss, P. O., Heller, C. H., & Fricke, K. J. 2000, *A&A*, **361**, 913
- Rand, R. J. 1992, *AJ*, **103**, 815
- Rizzi, L., Tully, R. B., Makarov, D., et al. 2007, *ApJ*, **661**, 815
- Rosolowsky, E. 2005, *PASP*, **117**, 1403
- Rozas, M., Beckman, J. E., & Knapen, J. H. 1996, *A&A*, **307**, 735
- Rozas, M., Zurita, A., & Beckman, J. E. 2000, *A&A*, **354**, 823
- Rozas, M., Zurita, A., Heller, C. H., & Beckman, J. E. 1999, *A&AS*, **135**, 145
- Saha, A., Thim, F., Tammann, G. A., Reindl, B., & Sandage, A. 2006, *ApJS*, **165**, 108
- Schaerer, D., & de Koter, A. 1997, *A&A*, **322**, 598
- Scoville, N. Z., Polletta, M., Ewald, S., et al. 2001, *AJ*, **122**, 3017
- Scowen, P. A., Dufour, R. J., & Hester, J. J. 1992, *AJ*, **104**, 92
- Smith, T. R., & Kennicutt, R. C., Jr. 1989, *PASP*, **101**, 649
- Terry, J. N., Paturel, G., & Ekholm, T. 2002, *A&A*, **393**, 57
- Thilker, D. A., Braun, R., & Walterbos, R. A. M. 2000, *AJ*, **120**, 3070
- Tully, R. B. 1988, *Nearby Galaxy Catalog* (Cambridge: Cambridge Univ. Press)
- Tully, R. B., Rizzi, L., Shaya, E. J., et al. 2009, *AJ*, **138**, 323
- van der Laan, T. P. R., Schinnerer, E., Emsellem, E., et al. 2013, *A&A*, **551**, A81
- Youngblood, A. J., & Hunter, D. A. 1999, *ApJ*, **519**, 55